ImPRINTS

- Allergens
- Cyt kines
- Immune and receptors
- Immunoglobulin superfamily
- CD
- Immunoglobulin related fold
- Other
- Virulence factors

Immunoglobulin superfamily

Many of the molecules being involved in the vertebrate immune responses share a common evolutionary precursor: the immunoglobulin homology unit. However, several other molecules with no known immunological functions have also been shown to share this same precursor element. Together, the genes encoding these related molecules have been defined as the immunoglobulin gene superfamily (IgGSF) and include both multigene and single-gene representatives. These IgGSF products represent an amazingly diverse array of functions from immune receptors to cartilage formation, reflecting the versatility if the shared common structure.

Members of the IgGSF have been defined by the presence of one or more regions homologous to the basic structural unit of immunoglobulin molecules, the Ig homology unit. The Ig homology unit provides the basic architecture upon which multiple sequence substitutions can be rapidly imposed, thus providing a diverse repertoire of recognition and target elements. It is characterised by a primary sequence about 70-110 residues in

length, with an essentially invariant disulphide bridge spanning 50-70 residues and several other relatively conserved residues involved in establishing a tertiary structure referred to as antibody fold. Two basic homology unit types have been defined from crystallographic analysis of the variable (V) and constant (C) regions of Ig. The tertiary structure of a V region is dominated by a series of nine antiparallel β strands, connected by variable-length loop sequences, that assume a characteristic barrel or sandwich like structure with two β sheets, stabilized by the disulphide bridge. There are four beta strands in one sheet and three on the other. The extra pair of β strands is essentially situated between the faces of the sandwich. The β strands are characterised by alternating hydrophobic and hydrophilic amino acid residues. The hydrophobic side chains are oriented toward the interior and help stabilize the interaction between the two sheets. The outpointing hydrophilic residues mediate the interaction. The disulphide bridge further stabilizes this basic structure, providing compact, globular domains that are relatively proteolytically insensitive.

The C region unit lack the pair of internal β strands, but otherwise assume the same general structure with a distinct, but overlapping series of conserved residues. The lack of this extra pair of strands decreases the distance between the two cysteine residues of C regions relative to those of V regions. The extra loop sequence connecting these two strands in V regions is critical to the formation of the antigen-binding pocket of antibodies. As new members of the IgGSF have been characterized, their homology units have generally been defined as either V- or C-like, based on primary sequence similarities and secondary structure predictions. Many of the more recently discovered members have a primary and secondary structural motif that, although shared among themselves, does not preferentially fit either the previously defined V or C homology units. It is a more compact unit, even shorter in length between the two cysteines than most C-like units. Accordingly, a new class of homology unit, denoted H, has also been defined to encompass these members. This third type of homology unit has also recently been defined as C2 (constant 2) mostly due to its lack of the V-defining extra b-strand pair. Overall, however, this unit appears to have a relatively equidistant relationship to both V and C homology units. Because both V and C units are each more closely similar

to H units than to each other, the H unit probably reflects a more primordial motif, suggesting that the original members of the superfamily arose in early metazoa and clearly carried out cell-surface recognition functions unrelated to vertebrate immunity.

Apart from the IgGSF members mediating specific recognition of antigens (mainly the Ig of B cells, the antigen-specific receptor of T cells, and the class I and II proteins of the MHC complex), at least 23 other distinct genes or gene families with no direct role in antigen interaction have also been identified as belonging to this superfamily. Most are single-gene members and most appear to encode distinct cell surface, receptor molecules. Although these genes are generally non-polymorphic, the diversity of the examples is striking. For convenience, these molecules can be loosely associated (not necessarily functionally or evolutionary) in eight categories:

1. Non-antigen presenting, β2-microglobulin-associated molecules:

   β2-microglobulin (the light chain of the MHC class I molecule) is a single C homology unit. It is probably divergently related to the MHC class II α chain and may be considered functionally an orphan MHC gene. It is encoded by a singe nonpolymorphic gene and is also found in association with the Qa and Tla MHC molecules and the CD1 family of antigens.

2. T-cell-associated molecules:

   Besides the TcR, T cells express a host of accessory molecules that are presumably involved in signal transduction, cell adhesion, and even the facilitation of antigen/MHC targeting. The CD4 and CD8 molecules are accessory molecules of T cells that appear to play an important role in facilitating T cell interaction with target cells.

3. Molecules expressed on both T cells and nervous system cells (e.g. the Thy-1 molecule, possessing a single V-like homology unit, found in abundance on

thymocytes and neurons as well as fibroblasts and a variety of other cells. It may possibly be involved in signal transduction).

4. Nervous-system-associated molecules (e.g. the N-CAM gene encodes five H-type N terminal homology units, a long connecting sequence, a transmembrane region, and a very large cytoplasmic domain. It is generally involved in cell-to cell interaction or adhesion in neuronal morphogenesis). Interestingly, from the number of genes expressed in both the brain and the immune system, the possibility of shared cell surface recognition functions, as well as the involvement of related molecules in some of the intriguing phenomena linking mental states and immune response, could be hypothesized.
5. Ig-binding molecules (e.g. the poly(Ig) receptor (p-IgR), whose function is to shuttle polymeric IgM and IgA antibodies from the blood side to the serosal side of mucous membranes. Its external portion is released during the process and is known as the secretory component-SC). The SC molecule has five V homology units. Comparison of the individual units indicates that they are each more closely related to each other than to other IgGSF members. Hence, they are likely to be the product of a series of internal duplication events that resulted in the expansion of a single unit sequence).
6. Growth factor/kinase receptors (e.g. PDGFR and CSF-1R, both possessing H-type homology units. It seem that the IgGSF receptor-kinase coupling took place through an exon shuffling event, and then they diverged to generate a family of growth factor receptors).
7. Miscellaneous and
8. Uncertain examples:

Several other sequences have been proposed as members of the IgGSF on the basis of limited stretches of identity with other members. Several of the polydomain members (i.e. CD4, PDGFR) have sequences that may be relic homology units. This possibility is supported by statistical analysis as well as their presence in molecules with other well-defined IgGSF domain sequences. Several of these relic regions do not have the conserved cysstein pair, However there is at least one excellent example of functional

immunoglobulin V region also lacking the disulfide bridge. Thus, it is clear that no single rule of membership in the IgGSF is absolute.